VOL. 16, NO. 21, NOVEMBER 2021 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

FORECASTING HOUSEHOLD ENERGY USAGE BASED ON EDGE DEEP LEARNING

Calveen Jon Elvin[1], Rozeha Abdul Rashid[1], A. S. Abdullah[1], Omar A. Aziz[1], M. A. Sarijari[1] and Teruji Ide[2]
[1]School of Electrical Engineering, Universiti Teknologi Malaysia, Skudai, Johor Bahru, Johor, Malaysia
[2]Department of Electrical and Electronic Engineering, National Institute of Technology, Kagoshima College, Japan
E-Mail: calveenjon@gmail.com

ABSTRACT

This study analyses the feasibility of using embedded processing device that will be used as edge deep learning device for real-time short-term load forecast for home energy management system (HEMS). We trained and implemented two well-known artificial neural network architecture, Multi-layer Perceptron (MLP) and Long Short-Term Memory (LSTM), into the Raspberry Pi as embedded computer and measure its prediction accuracy and processing time. The results revealed that MLP performs better than LSTM in terms of prediction accuracy and both models able to complete the forecasting process in less than a second for one-minute interval forecast. The embedded computing device has shown its potential in home energy management system field for reliable short-term load forecast which can significantly reduce the cost of the system for machine learning computation and thus reduce reliance on cloud computer for artificial intelligence inference.

Keywords: energy forecasting, edge artificial intelligence, deep learning.

1. INTRODUCTION

The innovation in green technology has gained a lot of attention recently to ensure the carbon dioxide emissions that cause global warming and climate change stay at a minimum level. Most developing countries still utilize fossil fuel such as coal for electricity generation. The rise in electricity demand can lead to increase in carbon emission as nowadays, electricity has become a part of human essentials. However, according to a study conducted by Lin B. *et al.*, the increase in electricity demands leads to negative greenhouse gas emission if clean energy is utilized but urbanization, population and industrialization lead to an increase in greenhouse gas emission instead (Lin and Li 2020). These relate to human behaviour towards the environment, especially in the consumption of natural resources. Harnessing solar energy has become a trend in some countries and various agencies have invested in the development of solar energy technology (Su and Wu 2020). The solar power system can be deployed easily at a lower cost and easy to maintain. This technology attracts a lot of consumers that live in a remote area who do not have direct access to electrical service.

Inefficient energy usage has been a major concern, especially in solar power system. The Home Energy Management System (HEMS) was introduced to help consumers in managing their energy consumption effectively by monitoring and controlling the load automatically (Lotfi, Abbou, and Abdi 2016). Such idea can motivate users to conserve the remaining stored energy for the next day usage or emergency needs as variation of weather and seasons can change hourly solar radiation patterns. The ability to predict changes in power consumption patterns at a specified time interval is essential for energy management system as it helps the system to plan on optimizing energy consumption.

Integrating Internet of Things (IoT) framework into HEMS has allowed the system to be deployed on a large scale as it able to communicate with other device or machine over a network, which creates opportunities for processing huge amount of data provided by sensing devices in the network. The abundance of data generated from HEMS has allowed integration with machine learning to become relevant in the energy management field. Machine learning has successfully improved HEMS accuracy in forecasting future power consumption based on changes in power demand patterns (Ahmad and Chen 2019; Ai, Chakravorty, and Rong 2019; Fallah *et al.* 2018). However, machine learning algorithms can easily exhaust HEMS computational resources and it can reduce HEMS performance in performing other tasks that require real-time operation such as power sensing and load controlling if such operations performed using a microcontroller. The problem involves limited computational resources can be mitigated by using cloud computing for performing a task that needs high computational power. Factors such as network latency and bandwidth should be considered as it can greatly impact the system response time. Edge computing has become significant in the IoT world as it brings computing resources close to IoT devices. Such integration can greatly reduce data transfer latency and bandwidth usage in data transferring over a network. Edge computing can assist limited computation resources IoT devices to perform computation intensive tasks such as machine vision and big data processing.

This study explores the use of an Artificial Neural Network (ANN) as a model for predicting household appliance power consumption using Raspberry Pi 3 Model B+. It is computationally expensive to predict power consumption using ANN. By deploying a low power micro-sized computer as an edge artificial intelligence device for performing prediction tasks, power

consumption can be predicted in real-time and data transferring latency between machines can be reduced without sacrificing HEMS response time and prediction model accuracy. This paper will be organized as follows. Section II reviews the previous works and compares their findings with this paper. Section III describes the methodology in detail and the results are presented in Section IV. Section V will be the conclusion of this paper.

2. RELATED WORKS

This section highlights the previous work related to load demand prediction using artificial neural network and deep learning edge computing.

The short-term load forecast is commonly used for controlling and scheduling the power system. One of the early attempts of using an Artificial Neural Network (ANN) for short-term load forecast was made by (Lee, Cha, and Park 1992) on forecasting one day ahead of hourly electric load using static and sequential approach. The method proposed using the backpropagation algorithm to estimate the weights of ANN and the author managed to obtain forecasting error percentage of 2% with computation time between 6.64 to 14.64 seconds using VAX 8550 computer. Another work by (Bobate and Ghate 2018) managed to achieve mean absolute percentage error (MAPE) of 6.398853% using the same backpropagation method for ANN weights estimate with only a single hidden layer in the ANN model. The author concluded that ANN is more efficient than the statistical method in forecasting load on special days. Liu *et al* (Liu *et al.* 2017) and Liang *et al* (Liang *et al.* 2019) proposed using Long Short Term Memory (LSTM) network for short-term load forecasting as it is capable to forecast at high accuracy. The work highlights the features of LSTM ability in handling information in arbitrarily long sequences and thus, the effect of past data results in efficient forecasting at high precision. Hybrid model using ANN and fuzzy logic for short-term load forecasting proposed by (Ma-WenXiao, Bai-XiaoMin, and Mu-LianShun 2002) to utilize the advantages of generalization capability and better at handling uncertain problems respectively. Besides the backpropagation method for estimating weights in ANN, there were some attempts on using evolutionary techniques such as method proposed by (Khan, Khan, and Ullah 2011) for optimizing the structure and weights of ANN using Cartesian genetic programming, however such methods can further complicate the structure of ANN which could make unnecessary structure expansion and complex neuron connections. Particle Swarm Optimization (PSO) algorithm for optimizing ANN was proposed by (Bashir and El-Hawary 2007) in accelerating the training process of ANN. However, the ANN weights optimization performance using PSO is greatly influenced by its parameters.

Edge computing contributes to IoT for allowing computation to be performed at the edge of the network. Due to various advancements in hardware, embedded devices may have plenty of computing resources for performing deep learning related tasks. A study done by Sakr *et al* (Sakr *et al*. 2016) implements a deep Convolutional Neural Network (CNN) and Support Vector Machine (SVM) to sort the waste accordingly by recognizing the type of waste using the image as an input. Based on their findings, the authors chose SVM into Raspberry Pi as edge devices for image classification as SVM gave higher prediction accuracy compared to CNN after trained them on a different machine. The average processing time for Raspberry Pi to perform image classification is 0.1 second with 0.005 seconds of standard deviation. The processing time is quite fast for Raspberry Pi considering image classification is a complex algorithm and it may consume a lot of computational resources. Another work made by Chen et. al (Chen *et al.* 2019) performing a novel crowd counting method using Speed-Switchable CNN (SsCNN) which can predict at high speed with high accuracy compared to conventional CNN. The model uses less computational power for image processing and the prediction time takes between 0.81 seconds to 4.06 seconds depending on the choice of speed and accuracy trade-off. The author has shown that low power embedded device capable of processing a deep learning algorithm makes it qualified as an edge computing device for crowd counting.

3. IMPLEMENTATION METHODOLOGY

A. Dataset Source and Preprocessing Method

The dataset used for training the prediction model is obtained from the UCI Machine Learning Repository named *Individual household electric power consumption Data set* (G. Hebrail 2012). This data set is a time-series data in which household power consumption for a single house in France is measured for more than 2 million minutes at one-minute interval. The data contains several attributes such as date, time, global active power, global reactive power and, etc. However, only global active power is considered for prediction modelling since active power is what we consume or utilized. Another dataset is obtained from (Rashid et al. 2019) that will be used for prediction benchmarking. It is important to provide two different datasets for model training and testing to eliminate the correlation between them.

The data set needs to be pre-processed before using it as an input for the prediction model. Missing or invalid values in a dataset can cause some issues during model training and prediction because it cannot process the data that does not represent a value or a real number. Therefore, any sections in the dataset that contains invalid value must be removed before any further dataset processing.

Normalizing the data is not mandatory in machine learning area but it is a good practice to normalize or scale the data as it may help to improve the model training convergence speed due to small initial errors made on the first iteration (Jin, Li, and Jin 2015). However, it does not affect the prediction accuracy of a model since raw and normalized data can achieve the same results despite their data range difference. In this paper, we use z-score normalization method to scale the data at zero mean. Equation (1) and (2) used to calculate the mean and

standard deviation respectively. The z-score is calculated using equation (3) for each data point.

$$\boldsymbol{\mu} = \frac{\mathbf{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{N}} \boldsymbol{x_i} \quad (1)$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \quad (2)$$

$$z_t = \frac{x_t - \mu}{\sigma} \quad (3)$$

The sliding window method is applied to the data once it has been normalized or scaled. In this paper, we propose using a window size of 10 data points as input data for the model to predict a single value. This will result the existence of correlation between the previous points to the current points.

B. Short-Term Load Forecast Model and Implementation

Machine learning algorithm used for this work is Artificial Neural Network (ANN). The ANN has the reputation of being a universal function approximator since it can solve most non-linear problems. The basic structure of ANN has three layers of nodes; input layer, hidden layer and output layer as illustrated in Figure-1. The input data fed to the input layer and the results from the computation of weighted sum of each neuron then parse it to another neuron in the next layer. Activation functions are applied to all layers except for input layer. A typical ANN can have multiple hidden layers that suits to the problems.

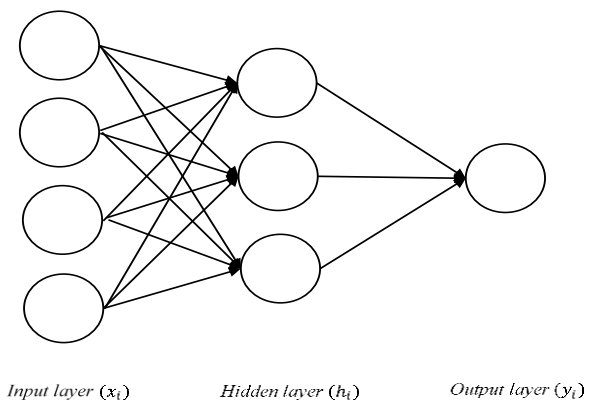

Figure-1. Artificial neural network structure.

A Multi-Layer Perceptron (MLP) is the most common types of neural network architecture used for various applications due its successes in predicting various types of regression and classification problems with high accuracy (Bharadwaj *et al.* 2018; Widiasari, Nugroho, and Widyawan 2018; Widyahastuti and Tjhin 2017; Zhai *et al.* 2016). Most applications that use MLP neural network benefit from multiple hidden layers instead of single hidden layers to form a complex model. The input layer of neural network forms a series of value that will be passed to each neuron in the next layer. The input of hidden layer neuron is the sum weighted from the input layer which the formula is as shown below:

$$h_i = \sum_{j=1}^{N} x_j w_{ij} \quad (4)$$

Activation function applied to sum weighted value obtains from input neurons and the result will be passed to the neuron in the next layer. The process of passing the information from one neuron to another neuron in the next layer is repeated until reaching the final layer which is the output layer. The equation of each output neuron in the output layer is the same as the equation (4) above and activation function applied after sum weighted has been computed.

Long Short-Term Memory (LSTM) is an artificial recurrent neural network architecture. It was proposed by Sepp Hochreiter and Jürgen Schmidhuber (Hochreiter and Schmidhuber 1997) to solve the vanishing gradient problems that occurred in vanilla recurrent neural network architecture by introducing memory cells to store information and gate units to regulate the information either to store or forget as illustrated in Figure 2. Both LSTM and vanilla recurrent neural network has hidden state which is used as looping mechanism. The recurrent neural network, especially LSTM architecture has an ability to retain memory of previous data for certain time makes it suitable for solving problems that involve time series prediction (Alhirmizy and Qader 2019; Ludwig 2019; Preeti, Bala, and Singh 2019). However, due to the complexity of the network structure, it requires lots of computational power for real-time prediction.¬

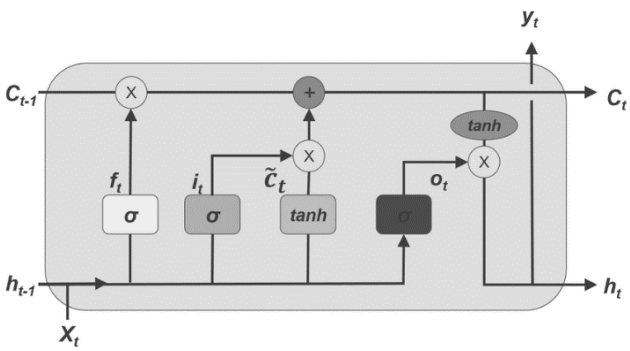

Figure-2. Long short-term memory cell structure.

This work proposes two different neural network models; MLP and LSTM architecture, with a total of four neuron layers to predict the future power consumption based previous consumption across the periods of time based the current power consumption patterns that are used as an input to the model. The number of neurons, hidden layers and activation function used are kept the same for all models. The performance of each model is measured to determine the suitable model for this work. The full details of the proposed models and training parameters are further described in Table-1.

Table-1. LSTM and MLP model specification and training parameters.

Items	Descriptions
Input layer	10 input neurons
Hidden layer 1 neuron(s)	10 neurons
Hidden layer 2 neuron(s)	10 neurons
Output layer neuron(s)	1 neuron
Activation function	Linear (hidden and output layer)
Training epochs	10
Forecast horizon	1
Training batch size	64

C. Forecast Model Assessment

Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are used to evaluate the performance of the model by measuring the differences between the actual and the predicted values. The RMSE and MAE equation are shown as below:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \widetilde{y_i})^2} \quad (5)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \widetilde{y_i}| \quad (6)$$

D. Forecast Model Testing Platform

Raspberry Pi is a single board computer used by many hobbyist, researcher and educator in developing an embedded project. There are several projects that utilize Raspberry Pi as part of the IoT framework as it has a full featured computer that embeds all necessary hardware and interfaces into a small circuit board.

This work proposes using Raspberry Pi 3 Model B+ as testing platform and its detailed hardware and software specifications are described in Table-2. The Raspberry Pi will be used for power consumption forecast but model training will take place on another platform with hardware that can handle model training activity due to high resource demand. The machine learning models are implemented in Python program using Keras framework. Keras is a high-level application programming interface for TensorFlow for developing and training deep learning models. It allows user to perform fast prototyping that concerns on the model design makes it beginner friendly. Since the model training takes place on the other machine, the trained model will be exported into files so that it can be easily transferred to Raspberry Pi by importing the model files into the program.

Table-2. Raspberry Pi 3 model B+ specifications.

Item	Description
System-on-Chip	Broadcom BCM2837B0, Cortex-A53 (ARMv8) 64-bit SoC
Central Processing Unit	1.4GHz 64-bit quad-core ARM Cortex-A53
RAM	1GB LPDDR2 SDRAM

4. EXPERIMENTAL RESULTS

In this section, the results of short-term load forecasting obtained from both MLP and LSTM are presented. Table 3 presents the prediction error and the prediction processing time for four different clusters. The dataset used for model evaluation [17] is sub-sampled at any section of the data set into four different clusters with a sample size of 100. The data clusters will be used as input data for the prediction model.

MLP model has the lowest MSE and RMSE values for all clusters compared to LSTM. Even though MLP does not have feedback loops in its structure for remembering past data, it can perform better than LSTM for time series data. In terms of processing time, MLP takes short time to predict the next value compared to LSTM since the MLP structure is a lot simpler compared to LSTM that has a lot of computation occurred in each neuron. In general, MLP is a lot more accurate compared to LSTM in forecasting the next power consumption. However, in terms of processing time, both models require less than a second to output the forecasted value which makes them suitable for real-time prediction task.

Both MLP and LSTM have a forecasting curve that follows the actual curve for all clusters. LSTM forecasting curve in all data clusters tends to predict the power consumption higher than the actual value. The MLP forecasting curve follows the actual curve with some minor error even at load spike. It can be seen also in Figures 3 and 5 the difference between LSTM and actual lines are larger compared to MLP with actual line in a condition where the power consumption remain constant. The gap between the predicted and the actual line reflects the RMSE and MSE values as shown in Table-3. Based on the results obtained, it is concluded that the MLP model is the most suitable for forecasting the household power consumption at a minute interval.

5. CONCLUSIONS

The problems in optimizing the household energy consumption have led to the development of a home energy management system. The embedded hardware that keeps evolving over the years has made implementing machine learning into the embedded device possible. This could greatly benefit the home energy management system for power consumption forecast which will promote effective household electrical load management. In this paper, MLP and LSTM neural network models are proposed for short-term load forecast. Based on the evaluation results, MLP model has the highest prediction

accuracy with the shortest processing time compared to LSTM. Therefore, MLP is the suitable model short-term load forecast. It is hoped that by unifying the MLP models with home energy management system, the system able to perform its task in optimizing the daily energy usage effectively without discomforting the user.

Table-3. MLP and LSTM prediction error and processing time.

ANN Models		Cluster 1	Cluster 2	Cluster 3	Cluster 4
MLP	**MAE**	**0.019**	**0.038**	**0.063**	**0.023**
	RMSE	0.026	0.073	0.143	0.029
	Processing Time (s)	0.188	0.189	0.191	0.200
LSTM	MAE	0.036	0.050	0.084	0.038
	RMSE	0.039	0.077	0.143	0.042
	Processing Time (s)	0.233	0.232	0.240	0.238

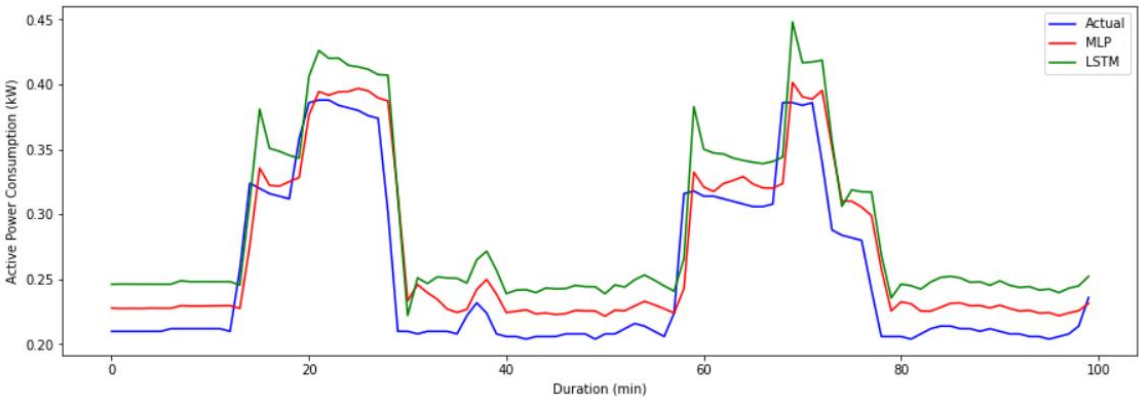

Figure-3. Power consumption forecast for cluster 1.

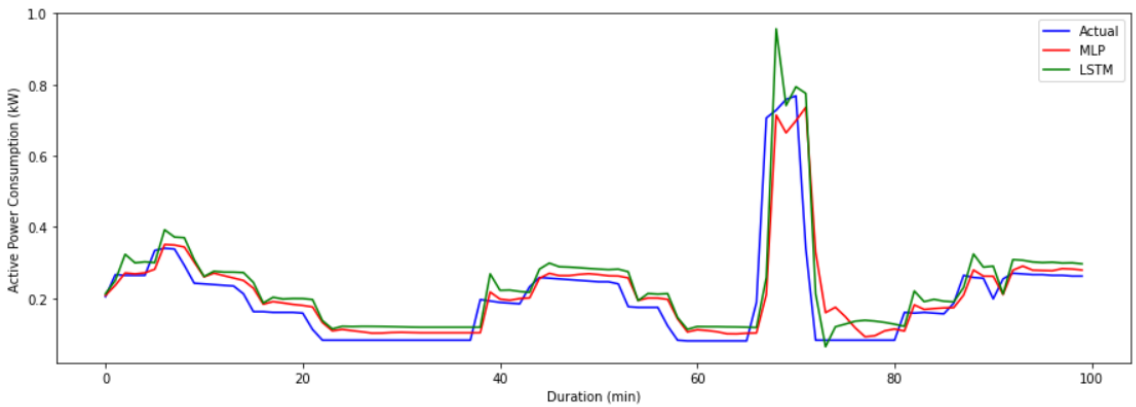

Figure-4. Power consumption forecast for cluster 2.

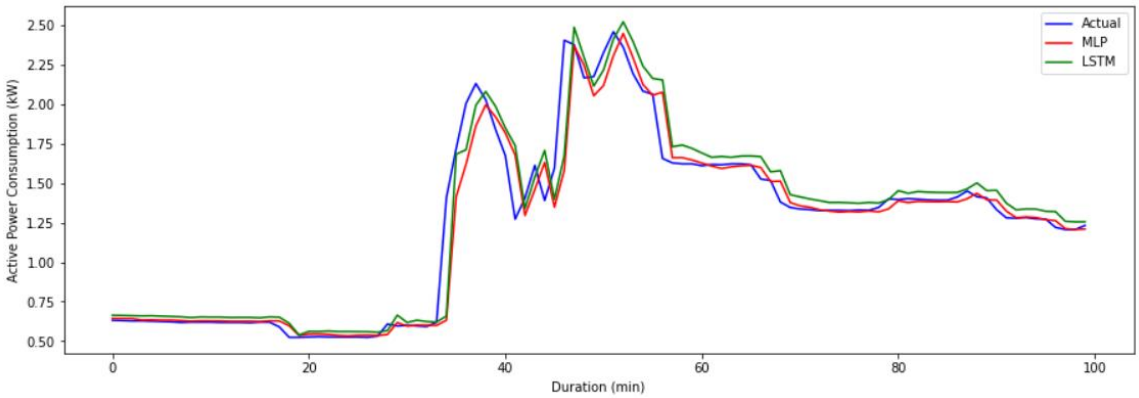

Figure-5. Power consumption forecast for cluster 3.

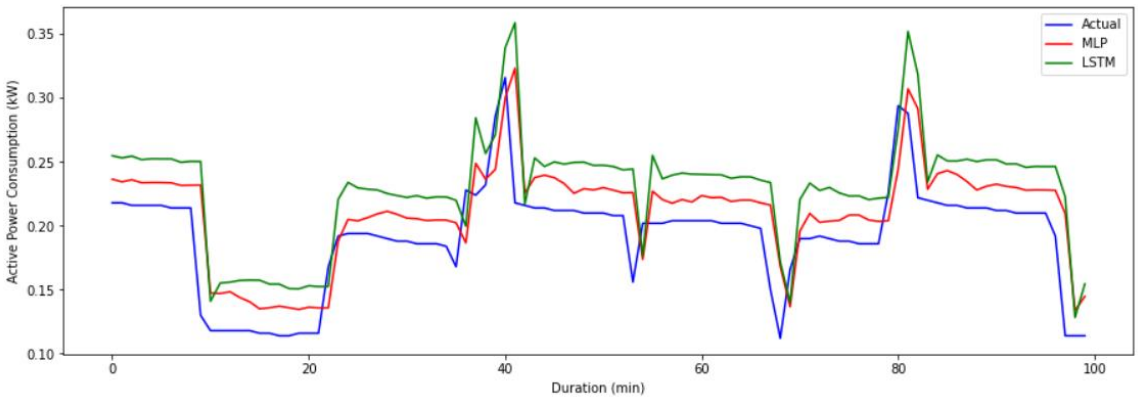

Figure-6. Power consumption forecast for cluster 4.

ACKNOWLEDGEMENT

The authors wish to express their gratitude to Ministry of Higher Education (MOHE), Malaysia and Research Management Center (RMC), Universiti Teknologi Malaysia for the financial support of this project under FRGS/1/2019/TK04/UTM/02/44 grant no: RJ130000.7851.5F174 titled Machine Learning based Interference Mitigation Scheme in Wireless Sensor Network for Internet-of-Things (IoT) Applications.

REFERENCES

Ahmad Tanveer and Huanxin Chen. 2019. Nonlinear Autoregressive and Random Forest Approaches to Forecasting Electricity Load for Utility Energy Management Systems. Sustainable Cities and Society. 45: 460-73.

Ai Songpu, Antorweep Chakravorty and Chunming Rong. 2019. Household Power Demand Prediction Using Evolutionary Ensemble Neural Network Pool with Multiple Network Structures. Sensors. 19(3): 721.

Alhirmizy Shaheen and Banaz Qader. 2019. Multivariate Time Series Forecasting with LSTM for Madrid, Spain Pollution. ICCISTA 2019 - IEEE International Conference on Computing and Information Science and Technology and Their Applications. 1-5.

Bashir Z. A. and M. E. El-Hawary. 2007. Short-Term Load Forecasting Using Artificial Neural Network Based on Particle Swarm Optimization Algorithm. pp. 272-75 in 2007 Canadian Conference on Electrical and Computer Engineering.

Bharadwaj Shrisha, B. S. Anil, Abhiraj Pahargarh, Adhiraj Pahargarh, P. S. Gowra and Sharath Kumar. 2018. Customer Churn Prediction in Mobile Networks Using Logistic Regression and Multilayer Perceptron (MLP). pp. 436-38 in Proceedings of the 2nd International Conference on Green Computing and Internet of Things, ICGCIoT 2018. IEEE.

Bobate P. D. and V. N. Ghate. 2018. Short-Term Load Forecasting with Artificial Neural Network. pp. 1276-80 in 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT).

Chen J., Q. Zhang, W. Zheng, and X. Xie. 2019. Efficient and Switchable CNN for Crowd Counting Based on Embedded Terminal. IEEE Access 7: 51533-41.

Fallah Seyedeh, Ravinesh Deo, Mohammad Shojafar, Mauro Conti and Shahaboddin Shamshirband. 2018. Computational Intelligence Approaches for Energy Load Forecasting in Smart Energy Management Grids: State of

the Art, Future Challenges, and Research Directions. Energies 11(3): 596.

G. Hebrail, A. Berard. 2012. Individual Household Electric Power Consumption Data Set. UCI Machine Learning Repository. Retrieved (http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption).

Hochreiter S. and J. Schmidhuber. 1997. Long Short-Term Memory. Neural Computation. 9(8):1735-80.

Jin Jian, Ming Li and Long Jin. 2015. "Data Normalization to Accelerate Training for Linear Neural Net to Predict Tropical Cyclone Tracks." Mathematical Problems in Engineering 2015.

Khan G. M., S. Khan and F. Ullah. 2011. Short-Term Daily Peak Load Forecasting Using Fast Learning Neural Network. pp. 843-48 in 2011 11th International Conference on Intelligent Systems Design and Applications.

Lee K. Y., Y. T. Cha and J. H. Park. 1992. Short-Term Load Forecasting Using an Artificial Neural Network. IEEE Transactions on Power Systems 7(1): 124-32.

Liang Fan, Austin Yu, William G. Hatcher, Wei Yu and Chao Lu. 2019. Deep Learning-Based Power Usage Forecast Modeling and Evaluation. Procedia Computer Science. 154: 102-8.

Lin Boqiang and Zheng Li. 2020. Is More Use of Electricity Leading to Less Carbon Emission Growth? An Analysis with a Panel Threshold Model. Energy Policy 137(October 2019): 111121.

Liu C., Z. Jin J. Gu and C. Qiu. 2017. Short-Term Load Forecasting Using a Long Short-Term Memory Network. - p. 1-6 in 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe).

Lotfi Jawad, Mohamed Fouad Abbou, and Farid Abdi. 2016. Overview of Home Energy Management Systems Architectures and Techniques. pp. 1-4 in Proceedings of 2015 IEEE International Renewable and Sustainable Energy Conference, IRSEC 2015. Marrakech, Morocco: IEEE.

Ludwig Simone A. 2019. Comparison of Time Series Approaches Applied to Greenhouse Gas Analysis: ANFIS, RNN, and LSTM. pp. 1-6 in IEEE International Conference on Fuzzy Systems. Vols. 2019-June. IEEE.

Ma-WenXiao, Bai-XiaoMin and Mu-LianShun. 2002. Short-Term Load Forecasting with Artificial Neural Network and Fuzzy Logic. pp. 1101-4 vol.2 in Proceedings. International Conference on Power System Technology. Vol. 2.

Preeti Rajni Bala and Ram Pal Singh. 2019. Financial and Non-Stationary Time Series Forecasting Using LSTM Recurrent Neural Network for Short and Long Horizon. pp. 1-7 in 2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019. IEEE.

Rashid R. A., L. Chin, M. A. Sarijari, R. Sudirman, and T. Ide. 2019. Machine Learning for Smart Energy Monitoring of Home Appliances Using IoT. pp. 66-71 in 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN).

Sakr G. E., M. Mokbel, A. Darwich, M. N. Khneisser and A. Hadi. 2016. Comparing Deep Learning and Support Vector Machines for Autonomous Waste Sorting. pp. 207-12 in 2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET).

Su Yi Huei and Jung Hua Wu. 2020. Using Synergy Approach to Investigate the Developmental Trends of Solar Energy. in IOP Conference Series: Materials Science and Engineering. Vol. 720.

Widiasari Indrastanti R., Lukito Edi Nugroho and Widyawan. 2018. Deep Learning Multilayer Perceptron (MLP) for Flood Prediction Model Using Wireless Sensor Network Based Hydrology Time Series Data Mining. pp. 1-5 in Proceedings - 2017 International Conference on Innovative and Creative Information Technology: Computational Intelligence and IoT, ICITech 2017. Vols. 2018-Janua.

Widyahastuti Febrianti and Viany Utami Tjhin. 2017. Predicting Students Performance in Final Examination Using Linear Regression and Multilayer Perceptron. pp. 188-92 in Proceedings - 2017 10th International Conference on Human System Interactions, HSI 2017.

Zhai Xiaojun, Amine Ait Si Ali, Abbes Amira and Faycal Bensaali. 2016. MLP Neural Network Based Gas Classification System on Zynq SoC. IEEE Access. 4: 8138-46.